New Media

Audio and Video on the Web

(Medical Education Web Page Series – Part 7)

The dramatic increase in the Internet bandwidth available to many users as a result of the proliferation of cable modem and DSL technologies has led to increasing interest in the use of audio and video technologies in Internet-based distance education. Audio and video delivered via the Internet can be used for a variety of educational purposes such as transmitting lectures and demonstrations. In particular, short video clips can be enormously beneficial in demonstrating clinical procedures such as intubation or line insertion. While video files often take up considerable space, audio files usually have much more modest space requirements, depending on the quality of recording delivered.

Multimedia files such as audio or video come in two general formats: “downloaded” format and “streaming” format. In the case of the former, the file must be completely received before it can be played, while in the case of streaming files, the material can be played before the material is fully received. Another difference is that downloaded files can usually be saved to hard drive for later use while some streaming formats do not allow the material to be permanently saved so that one can replay them at will. (This “single play only” feature appeals to many music distributors who fear that their material could be spread over the Internet without any restriction such as happened when Napster first set up shop).

Sound, image and video files can be uncompressed or compressed, but compression is almost always used when multimedia materials are transmitted via the Web. This is necessary in order to reduce transmission time. Compression can be as much as 90% or more, depending on the technology used. Furthermore, compression can be lossless or lossy, depending on whether or not the original data can be mathematically reconstructed exactly by a reverse operation.

Incorporation of multimedia into Web pages is now relatively easy when advanced Web page production packages such as Microsoft Front Page are used. Such

packages often allow incorporation of multimedia files using simple “click and drag” methods. This shifts the challenge to that of producing good quality content. A variety of inexpensive audio and video editing tools are available for that purpose.

For more information, please consult the information shown in Tables I and II. To see some sample Web-based videos on a variety of anesthesia topics, visit <http://gasnet.org/videos/index.php> (you will need to download the RealVideo G2 player before you can view these videos).

D. John Doyle MD PhD FRCPC
Cleveland, Ohio

TABLE I Glossary of Audio/Video Terms

*Modified from materials made available at
<http://www.columbia.edu/itc/itc/webdev/design/av-intro.html> and elsewhere*

Aspect Ratio refers to the ratio of a video image’s width to its length. Video seen on regular television has an aspect ratio of 4:3 while widescreen video seen in “letterbox format” movies and DVD video employs an aspect ratio of 16:9. The “DV” format used in digital video cameras uses an aspect ratio of 3:2 (with a frame size of 720 × 480).

AU is a standard for Unix/Linux computers that gained widespread use for sounds on the Internet a few years ago when the World Wide Web first emerged. It is used less commonly now.

AVI; Audio Video Interleave. AVI is a video format developed by Microsoft and is one of the more common formats for PC video data. The format is interleaved such that video and audio data are stored consecutively (i.e., a segment of video data is immediately followed by a segment of audio data).

Colour depth (also known as *bit depth* or *pixel depth*) is a term that applies to both video and static images and refers to the maximum possible number of colours. A 24-bit colour depth scheme allows for 16.7 million distinct colours in an image. Achieving higher bit depth in video requires larger file sizes, but generally provides superior image quality.

Codec is short for “compressor/decompressor”. A codec is any technology for compressing and decompressing digital data, although the term usually applies to audio and video signals. The rationale for the use of a codec is that the digital representation of audio and video signals requires large amounts of data, often too much to send efficiently over an ordinary 56 Kbps modem connection. Digital media presented over the web is therefore often compressed to a much smaller size by selectively removing unim-

portant data from the file. The codec is thus the method (algorithm) used to compress and later decompress the media. Codecs can be implemented in software, hardware, or a combination of both. Some popular codecs for computer video include MPEG, Indeo and Cinepak.

Data Compression. Storing data in a format that requires less space than usual. Data compression is particularly useful in multimedia and in telecommunications because it enables devices to transmit the same amount of data using fewer bits. There are a variety of data compression techniques, of which the MP3 format is a particularly popular one for audio data.

Data rate refers to the transmission rate of data. Most modems top out at around 56 kilobits per second (Kbps), or 7 kilobytes per second (KBps) while connections using a Local Area Network (LAN) can be much higher. Data rate capability is an all-important factor upon which many other variables depend. These variables include colour depth, codec technique, frame rate, sampling rate, sampling resolution, and video size. Uncompressed NTSC video (the usual format used in North America) takes up about 27 megabytes per second, while compressed DV video takes up only about 3 megabytes per second.

Frame rate refers to the number of frames per second (fps) in a video. A smaller frame rate decreases the data rate, but makes the video appear jerky. Full motion video runs 25 to 30 fps. An acceptable rate for web video is 10 to 15 fps, especially if the action is relatively slow, which is the case in most medical procedure videos.

JPEG. Short for Joint Photographic Experts Group, JPEG is a "lossy" compression technique for colour images. Although it can reduce file sizes to as little as 5% of their precompression size, some detail is lost in the compression.

Keyframe refers to a "reference frame" in a video. A video codec works by taking periodic snapshots of a video, and redraws that snapshot as the video progresses. Only changes in sequential frames are redrawn, giving the appearance of motion. Using less frequent keyframes reduces the data rate, but also reduces the quality of the video.

Multimedia. The use of computers to present text, graphics, video, animation, and sound in an integrated way.

MP3 (MPEG-1 Layer-3 audio) - a popular audio format with near CD quality sound at a compression ratio about 10:1.

MPEG. Short for Moving Picture Experts Group, a working group of ISO. The term refers to the family of digital video compression standards. MPEG achieves a high compression rate by storing only the changes from one frame to another, instead of each entire frame. There are three major MPEG standards: MPEG-1, MPEG-2 and MPEG-4.

- The most common implementations of the MPEG-1 standard provide a video resolution of 352-by-240 at 30 frames per sec (fps). This produces video quality slightly below the quality of conventional VCR videos.
- MPEG-2 offers resolutions of 720 x 480 and 1280 x 720 at 60 fps, with full CD-quality audio. This is sufficient for all the major TV standards, including NTSC. MPEG-2 is used by DVDs.
- MPEG-4 is a graphics and video compression algorithm standard that is based on MPEG-1, MPEG-2 and Apple QuickTime technology. Wavelet-based MPEG-4 files are smaller than JPEG or QuickTime files, and are designed to transmit video and images over a narrower bandwidth.

Ripper. A software program that "grabs" digital audio from a CD and transfers it to a computer's hard drive, usually in MP3 format. DVD rippers are also commonly available but may not be legal.

Sampling rate refers to the number of samples per second, and is a determinant of audio quality. It is the audio equivalent of frame rate. Sampling rates for audio signals typically range from 8000 Hz to 48000 Hz or more. CD-quality audio is 44100 Hz at 16 bits resolution.

Sampling resolution refers to the precision with which a sound is recorded, usually either 8 or 16-bit. CD-quality audio is 16-bit.

SMIL. Synchronized Multimedia Integration Language. (SMIL, pronounced "smile"). This technology enables simple authoring of interactive audiovisual presentations. SMIL is typically used for "rich media"/multimedia presentations which integrate streaming audio and video with images, text or any other media type. SMIL is an easy-to-learn HTML-like language, and many SMIL presentations are written using a simple text-editor. For more information visit <http://www.w3.org/AudioVideo/>

Streaming is a method of delivery method of audio and video media whereby the material will play as soon as enough information arrives at the client; the full file need not be downloaded. By contrast, non-streaming audio and video require the whole file to be downloaded before playback. For audio data, the de facto Web standard is *RealAudio* from RealNetworks (www.real.com).

WAV. An early nonstreaming format for sound developed jointly by Microsoft and IBM. Support for WAV files was built into Windows 95 making it the de facto standard for sound on PCs. WAV sound files end with a .wav extension and can be played by nearly all Windows applications that support sound.

TABLE II Selected Audio/Video Links

http://www.real.com
http://www.realnetworks.com
http://www.apple.com/quicktime
http://www.windowsmedia.com
http://www.mpeg.org
http://www.w3.org/AudioVideo/
http://www.xingtech.com/
http://streamingmediaworld.com/
http://builder.cnet.com/webbuilding/pages/Graphics/StreamingMedia/